



*Biology and Philosophy* 00: 1–17, 2005.

© 2005 Kluwer Academic Publishers. Printed in the Netherlands.

## 1 Parsimony and the Fisher–Wright debate

2 ANYA PLUTYNSKI

3 *University of Utah, 260 Central Campus Drive, Room 341 OSH, Salt Lake City, UT 84112, USA*

4 (*e-mail: plutynski@philosophy.utah.edu*)

5 Received 26 August 2003; accepted in revised form 1 December 2004

6 **Key words:** Bayes' theorem, Density dependence, Epistasis, Genetic drift, Likelihood, Parsimony,  
7 Probability, Ronald Fisher, Shifting balance, Sewall Wright

8 **Abstract.** In the past five years, there have been a series of papers in the journal *Evolution* debating  
9 the relative significance of two theories of evolution, a neo-Fisherian and a neo-Wrightian theory,  
10 where the neo-Fisherians make explicit appeal to parsimony. My aim in this paper is to determine  
11 how we can make sense of such an appeal. One interpretation of parsimony takes it that a theory  
12 that contains fewer entities or processes, (however we demarcate these) is more parsimonious. On  
13 the account that I defend here, parsimony is a 'local' virtue. Scientists' appeals to parsimony are not  
14 necessarily an appeal to a theory's simplicity in the sense of it's positing fewer mechanisms. Rather,  
15 parsimony may be proxy for greater probability or likelihood. I argue that the neo-Fisherians  
16 appeal is best understood on this interpretation. And indeed, if we interpret parsimony as either  
17 prior probability or likelihood, then we can make better sense of Coyne et al. argument that  
18 Wright's three phase process operates relatively infrequently.

19

## 20 Introduction

21 In 1914, Sturteyvant defended the view that linkage is best explained by the  
22 hypothesis that genes are aligned on a chromosome, as opposed to a view  
23 favored by Bateson and Castle, the 'reduplication' hypothesis.<sup>1</sup> Summing up  
24 his rationale, he wrote that 'the chief advantage of the chromosome hypothesis  
25 of linkage... *seems to be its simplicity.*' In retrospect, we know that Sturteyvant  
26 was correct, but in this passage, his reasoning seems spurious. What does  
27 Sturteyvant mean when he suggests that the chromosomal hypothesis is 'sim-  
28 pler'? Why does the greater simplicity of a hypothesis count as a reason that  
29 favors it (perhaps defeasably) over its rival? Carlson, in his discussion of  
30 Sturteyvan'ts rationale, is skeptical of the virtue of simplicity as grounds for  
31 preferring one hypothesis to another, calling such appeals merely 'aesthetic'  
32 (Carlson 1966). Many philosophers have expressed similar skepticism.

33 This skepticism is often connected to a more general skepticism about the  
34 simplicity of nature. Why assume that the hypothesis that admits the fewest  
35 causes, entities, or processes, (however these are demarcated, and, all else being  
36 equal) is the more likely to be true? It would seem that such a view presupposes

<sup>1</sup> For details on the reduplication hypothesis, see Carlson, *The Gene: A Critical History*, 1966,  
p. 56. (The details of the theory are not essential to understanding this example.)

an ontological thesis. According to this thesis, we are to assume that nature is in fact simple, or that the natural world is governed by simple laws, a few fundamental forces, or, as Newton suggested, 'Nature does nothing in vain, and more is in vain where less will serve' (Newton 1729). In other words, if we assume that nature takes the shortest way, in every case where we are assessing competing explanations for some set of phenomena, we ought to choose that explanation which invokes fewer rather than more causes, entities, or processes.

But such reasoning seems spurious. If one understands simplicity, or, parsimony<sup>2</sup> as economy of process, it is tempting to dismiss appeals to parsimony by scientists, and especially appeals made by biologists. There seems to be no principled reason why one ought to assume that the biological world is simple, or that an explanation of that world which makes appeal to fewer entities or processes is better. Indeed, in the biological realm, there seem to be good reasons not to expect that nature will take 'shortest way.' Jacob (1977) and Gould and Lewontin (1979) have taught us well that 'nature is a tinkerer'; phylogenetic inertia, developmental constraints, chance, and the possibility of multiple solutions to the same adaptive problems are all at play in yielding extant morphology, distributions of traits in populations, and biogeographical distributions of species. At best, appeal to simplicity in this global sense is a heuristic virtue, insofar as simpler hypotheses are perhaps easier to test, on a some contrals of simplicity. But the parsimony of some thesis in this respect does not lend it greater credibility. The global ontological thesis seems hardly warranted in the case of biology. Sturtevant is clearly unwarranted in taking simplicity *per se* as a reason to prefer the chromosomal theory.

However, is this what Sturtevant meant to say when he argued that the chief virtue of the chromosomal theory its simplicity? In Sturtevant's view, according to the alternative, 'reduplication' hypothesis:

65

We are forced to assume an enormously complex series of cell divisions, many of them differential, proceeding with mathematical regularity and precision, but in a manner for which direct observation furnishes no basis. It seems to me that it is not desirable to assume such a complex series of events unless we have extremely strong reasons for doing so. I can see no sound reason for adopting the reduplication hypothesis. It *apparently rests on two discreditable hypotheses: somatic segregation, and the occurrence of members of the 3:1, 7:1, 15:1, etc., series of gametic ratios in more cases than would be expected from a chance distribution...*

<sup>2</sup> Twardy (personal communication) raises the question of whether simplicity and parsimony are one and the same virtue. Though they are often used interchangeably, my claim in this paper is that parsimony, or rather, what scientists often mean by parsimony, is not simplicity, in the sense of fewer parameters in one's model or fewer mechanisms or processes invoked. So, I agree with Twardy that the two virtues may be pulled apart.

75 76 ... the chief advantage of the chromosome hypothesis of linkage... seems  
 77 to be its simplicity. In addition, it *appeals to a known mechanism*... It  
 78 explains everything that any of the forms of the reduplication hypothesis  
 79 does and in addition *offers a simple mechanical explanation for the fact*  
 80 *that 'secondary series' are always smaller than Trow's special hypothesis*  
 81 *calls for them to be*. On the reduplication hypothesis this fact must merely  
 82 be accepted, for, I think, it can not be explained (Sturteyvant, in Carlson  
 83 1966).

84 Once we examine Sturteyvant's argument in closer detail, we can see that he  
 85 is not suggesting that we ought to prefer the chromosomal hypothesis solely on  
 86 the grounds that it has a greater economy of process. Moreover, the argument  
 87 does not appear so obviously misguided. Indeed, once one examines Sturtey-  
 88 vant's reasoning behind the claim that the 'simpler' theory is preferable, the  
 89 argument begins to look quite persuasive. Sturteyvant claims, first, that on the  
 90 alternative view, we have to posit events that we have no independent evidence  
 91 for. Second, he claims that the alternative view 'rests on' or presupposes  
 92 hypotheses that are 'discreditable.' Third, his preferred chromosomal  
 93 hypothesis makes use of a mechanism whose operation is well understood.  
 94 Fourth and finally, there are phenomena that the alternative hypothesis cannot  
 95 explain and the chromosomal hypothesis can. Thus, in Sturteyvant's view, we  
 96 ought to prefer the chromosomal hypothesis because it operates via a known  
 97 mechanism, does not presuppose questionable hypotheses, and can explain  
 98 phenomena that the alternative hypothesis leaves mysterious. In other words,  
 99 the chromosomal hypothesis is 'simpler,' by which he means, makes better  
 100 sense of the evidence to hand.

101 I wish to suggest that appeals to simplicity or parsimony in the biological  
 102 context are often shorthand for more elaborate and well-grounded rationales,  
 103 once one unpacks the argument carefully. This is the case in the paper by Coyne  
 104 et al. where they claim that mass selection is a more parsimonious hypothesis than  
 105 Wright's three phase shifting balance process. Thus, my interpretation is at odds  
 106 with that of Skipper, who argues (Skipper 2002) that we ought to interpret the  
 107 neo-Fisherian's appeals to parsimony in the same vein as I originally interpreted  
 108 Sturteyvant above. In other words, Skipper argues that by 'parsimony', the neo-  
 109 Fisherians meant 'economy of process.' His summary of the neo-Fisherian po-  
 110 sition is as follows, 'If the evolution of populations can be explained adequately  
 111 via a theory that postulates a small economy of entities and processes, there is no  
 112 need to invoke a theory with a larger economy of entities and processes' (p. 360).  
 113 Skipper contends that this view is an instance of a naïve appeal to parsimony, one  
 114 which sacrifices realism and precision for generality. The neo-Fisherians, he ar-  
 115 gues, ascribe to a 'Newtonian' ideal, according to which there is one theory (what  
 116 Skipper calls Fisher's 'large size theory'), that explains all of the phenomena in  
 117 some domain. 'However,' Skipper warns us, 'considerable care must be taken in  
 118 drawing a close connection between explanatory adequacy, generality, and

119 parsimony because explanatory adequacy need not be so closely connected with  
120 generality' (p. 361).

121 I agree with Skipper that we should certainly take care in assuming that there is  
122 a close connection between explanatory adequacy, generality and parsimony. In  
123 other words, a more explanatory theory is not necessarily a more general one  
124 (though, this will depend in part upon the request for explanation, or, the  
125 pragmatic dimensions of the why-question). Moreover, I agree heartily with  
126 Skipper that the theory which posits fewer mechanisms, entities, or processes  
127 (however we count these), all else being equal, is not necessarily more likely to be  
128 true. However, this is not what the neo-Fisherians were suggesting. Just as in the  
129 case of the example of Sturtevant's argument for the chromosomal theory, once  
130 we examine the neo-Fisherian's appeal to parsimony more closely, it is not so  
131 obviously misguided. In claiming that mass selection is 'more parsimonious,'  
132 Coyne et al. (1997) are claiming not that it is more general or more simple in the  
133 sense of invoking fewer entities or processes. Rather, by 'more parsimonious'  
134 they mean that mass selection operates via a known mechanism which is  
135 empirically well-established, does not depend upon presuppositions that are  
136 questionable, and finally, that the evidence tells against the alternative hypothesized  
137 mechanism as operating very frequently.

138 Moreover, their argument is not that the Fisherian model is more explanatory  
139 simply because it is more general. In those cases where the Wrightian three phase  
140 process is occurring, the Wrightian model would certainly be the best explanation.  
141 However, they claim that there are good empirical and theoretical grounds  
142 for these cases being rather rare. And thus, we should expect more requests for  
143 explanation of this or that adaptation to be satisfied by the Fisherian model. In  
144 other words, it does not explain more because it is more general; rather it is more  
145 general because it explains more.

146 Before I defend this thesis, I wish to make clear what I am not attempting in this  
147 paper. My aim is not to provide an overview of the debate between the neo-  
148 Wrightians and the neo-Fisherians. Rather, my aim is to clarify what is meant by  
149 'parsimony' by Coyne et al. and in the process, to defend an account of parsimony  
150 that takes it to be a genuinely epistemic, and not merely aesthetic or pragmatic  
151 virtue. I take my cue from Sober, in a paper titled 'Let's Razor Ockham's Razor'  
152 (1990) and *Reconstructing the Past* (1988). He claims that by giving close attention  
153 the specific context in which appeals to parsimony are made by scientists, one  
154 may come to understand that these appeals may sometimes be understood as  
155 appeals to either higher prior probabilities or higher likelihood. Sober's model  
156 for unpacking appeals to simplicity employs Bayes' Theorem.<sup>3</sup>

<sup>3</sup> The model is 'Bayesian' in the sense that it uses Bayes' theorem, an uncontroversial theorem in the probability calculus. It does not commit one to a particular view on reconditionialization, subjective probabilities, etc.. I.e., one does not need to subscribe to Bayesianism (whatever that means; reasonable people disagree) in order to see why Bayes's rule is a useful way to elucidate scientific inference in cases such as these involving competing hypotheses. I wish to emphasize that I am not committed here to any theses about whether or not probabilities are best understood as degrees of belief.

157 *Fisher vs. Wright*

158 R.A. Fisher and Sewall Wright were two theoretical population geneticists  
 159 working in the early twentieth century who placed different emphasis on dif-  
 160 ferent factors in the evolutionary process. According to Wright, selection alone  
 161 is not sufficient to generate adaptive novelty. By 'novelty' here, I mean to  
 162 emphasize that according to Wright, selection alone could not suffice for major  
 163 adaptive changes due to radical changes in the genetic constitution of a pop-  
 164 ulation. Wright did not deny that selection could yield adaptative evolution.  
 165 However, he thought that there must be a 'balance' of 'forces' at play in  
 166 evolution (selection, drift, etc.), or a population would eventually become  
 167 'stuck' atop adaptive peaks. This is because selection acts on genes in combi-  
 168 nation, and, according to Wright, there is pervasive epistasis for fitness (or,  
 169 breaking up such gene combinations may be maladaptive). Unless a population  
 170 is subdivided, novel adaptive gene combinations will not come about, and  
 171 populations of organisms may become 'stuck' atop suboptimal peaks in the  
 172 'adaptive landscape.'

173 The adaptive landscape is a model of the relative fitness of different gene  
 174 combinations, with the horizontal axes representing different genotypes, and  
 175 the vertical axis representing relative fitness. For Wright, the problem of how  
 176 to move from suboptimal to higher 'adaptive peaks' in the field of gene com-  
 177 binations was *the* key problem that theoretical biologists must solve, for the  
 178 reasons I stated above. Wright wrote, 'The problem of evolution as I see it is  
 179 that of a mechanism by which the species may continually find its way from  
 180 lower to higher peaks in such a field' (1932). In other words, Wright's question  
 181 was, 'How could one shift from one "balanced" gene combination to another,  
 182 across what must be deep valleys of low fitness?' According to Wright, the most  
 183 effective means of traversing such peaks is via a three phase 'shifting balance'  
 184 process of isolation of small subpopulations, intrademic (within group) and  
 185 interademic (between group) selection. Wright believed this three phase process  
 186 to be the main means of generating adaptation, and perhaps also, many cases  
 187 of speciation.

188 If the reader has already begun to worry about the epistemic or ontological  
 189 status of adaptive landscapes, then she (or he) is not alone. Indeed, there has  
 190 been a flurry of papers in the evolution literature recently about the status and  
 191 shape of the 'adaptive landscape.'<sup>4</sup> Philosophers like Ruse (1993) have ex-  
 192 pressed skepticism about the way in which the adaptive landscape metaphor  
 193 misleads scientists. Fisher (letter to Wright, in Provine 1986 ■Au: AQ: Citation  
 194 Provine (1986) is not in list please add to list.■) was one of the first to raise  
 195 concerns about the landscape. Was it indeed three dimensional? Or, as we  
 196 consider a greater and greater number of genes in combination, could the  
 197 landscape that describes relative fitness of different genotypes become multi-  
 198 dimensional? Is the landscape 'rugged' in the way Wright suggested? Are the

<sup>4</sup> See especially, Gavrillets (1996, 1997).

199 'peaks' static, or could there be 'ridges' arising between peaks over time, either  
 200 due to assortative mating, or changes in the environment, such that epistasis  
 201 for fitness is not necessarily a barrier to adaptive evolution and genuine evo-  
 202 lutionary novelty? These are exactly the worries that Coyne et al. the modern  
 203 neo-Fisherians have raised. Some might argue that one is stacking the deck  
 204 against Wright by arguing that by adopting the adaptive landscape metaphor,  
 205 his model makes presuppositions which are questionable. Surely we should first  
 206 look at the empirical evidence for or against those presuppositions? Indeed,  
 207 this is what Coyne et al. (1997) do at length.

208 As mentioned above, Fisher did not agree with Wright's presuppositions  
 209 about either the extent of genetic epistasis for fitness, nor the metaphorical  
 210 adaptive landscape which depends upon this assumption. Fisher suspected that  
 211 the landscape was not three dimensional, but rather multidimensional, such  
 212 that adaptive evolution could occur along any of several trajectories of gene  
 213 frequency change. Epistasis, according to Fisher, was not so pervasive that  
 214 populations must become 'stuck' atop adaptive peaks. Moreover, Fisher  
 215 pointed out that while populations may well be structured in nature, migration  
 216 every generation between groups makes questionable the effectiveness of drift  
 217 as a way isolating novel gene combinations and thus of 'peak-shifting'. This is  
 218 not to say that Fisher thought that drift did not operate in evolution. Rather,  
 219 Fisher thought it unlikely that drift played an important role in *adaptive*  
 220 evolution. Fisher also infamously believed that one could treat effective pop-  
 221 ulation size as the entire breeding population, which he assumed to be quite  
 222 large (on the order of infinity!). And since selection is more effective than drift  
 223 in populations of large size (where  $4Ns > 1$  ( $N$  = population size,  
 224  $s$  = selection coefficient), selection must be the main factor in adaptive evo-  
 225 lution.

226 In 1997, Coyne et al. assessed the empirical and theoretical support for the  
 227 Wrightian vs. Fisherian model in the journal *Evolution*. They concluded that  
 228 there was relatively little evidence that *Wright's particular three-phase process*  
 229 plays a significant role in the evolution of adaptations. They were not  
 230 questioning the fact of Wright's influence and contributions, or that one or  
 231 another phase of the process might be important in adaptive change (e.g.  
 232 group selection), *only* that the three phase shifting balance process in par-  
 233 ticular was a major mode of adaptive change. Moreover, they were not  
 234 suggesting (with Fisher), that the effective population size of most popula-  
 235 tions was the entire breeding population, or that this was on the order of  
 236 infinity. Rather, they drew the more modest conclusion that selection ought  
 237 to be preferred as the more 'parsimonious' explanation for adaptation over  
 238 the three phase shifting balance model. In reply, Wade and Goodnight defend  
 239 Wright, and attacked the neo-Fisherians for their naïve commitment to  
 240 parsimony as a theoretical virtue. Lewontin has argued that biology is an  
 241 'epistemologists' paradise' exactly because of arguments of this sort; we have  
 242 two theories, both seemingly able to account for the same phenomena: which  
 243 do we choose?

## 244 Parsimony: not what you thought it was

245 At first glance, it may seem plausible that when Coyne et al. claim that the  
 246 Fisherian view is more parsimonious, they mean that it has the greatest  
 247 economy of process. Fisher's mass selection is 'less complicated' than Wright's  
 248 three phase process because one mechanism (selection) is operating as opposed  
 249 to four or five, depending upon how you count (isolation, drift, intra- and  
 250 interdemec selection, migration). I wish to counter this interpretation. The view  
 251 that they defend is not that Fisher's theory is preferable because it involves  
 252 fewer processes. Rather, it is that the Fisherian process has a higher prior and  
 253 likelihood than Wright's three phase process. More specifically, their claim is  
 254 that given the body of evidence at hand, the chance of shifting balance playing  
 255 a significant role in adaptive evolution is low relative to the alternative, because  
 256 the conditions required for it to operate are not very likely to be met in nature.

257 In short, the neo-Fisherian model

258 \*Operates via known mechanisms

259 \*Does not depend on questionable presuppositions

260 [i.e. uses presuppositions we accord a higher prior, hence has a higher prior]

261 \*Fits the data better than Wright.

262 [Hence higher likelihood.] In short, it has a higher likelihood and prior than  
 263 the alternative shifting balance theory.

## 264 The Bayesian way

265 In a (1990) essay, and again in his (1988) book, Elliot Sober gives careful  
 266 scrutiny to the notion of parsimony.<sup>5</sup> Below, I will sketch Sober's analysis of  
 267 parsimony and suggest that this is one useful way to understand what is meant  
 268 by 'parsimony' in the Fisher-Wright debate. As Sober himself points out, one  
 269 need not subscribe to Bayesianism in order to see the use of Bayes's rule as one  
 270 way of making sense of how scientists revise their views.

271 First, some terminological distinctions. Sober points out that Bayes's theo-  
 272 rem provides a useful way of characterizing the considerations that might affect  
 273 one's assessment of a hypothesis' plausibility. The theorem says that the con-  
 274 ditional probability of some hypothesis  $H$  ( $\Pr(H/e)$ ) is:

$$\Pr(e/H)\Pr(H)]/\Pr(e)$$

276 So, when comparing two hypotheses,  $H1$  and  $H2$ , their posterior probability is  
 277 influenced by two factors: their priors and their likelihoods:

$$\text{Or, } \Pr(H1/e) > \Pr(H2/e) \text{ iff } \Pr(e/H1)\Pr(H1) > \Pr(e/H2)\Pr(H2)$$

279 What does all this mean?

<sup>5</sup> Though, more recently, (Sober and Wilson 1998) Sober and Wilson demarcate several different forms of appeal to parsimony by Williams (1966), ■Au: AQ: Citation Williams (1966) is not in list please add to list. ■ not all of which are epistemic.

280  $\Pr(H1)$  is the prior probability of  $H$ .  $\Pr(e/H1)$  is  $H1$ 's likelihood. What is  
 281 likelihood? For one theory to have a higher likelihood, roughly, the evidence  
 282 will be more likely obtain than on the alternative theory. Likelihood is not the  
 283 same, however, as explanatory power. For example, if  $H1$  is 'rain tomorrow'  
 284 and  $e$  is 'today's barometric reading is 29 torr,' then,  $H1$  would have a high  
 285 likelihood relative to the evidence, but it would not explain the evidence.

286 One way of thinking about what scientists do when they're assessing the  
 287 explanatoriness of two competing hypotheses is that they're evaluating their  
 288 relative likelihoods, and/or antecedent plausibility. According to Sober, the  
 289 appeal to parsimony in such contexts are not necessarily appeals to greater  
 290 economy of process (though he does not rule out that they could be, given  
 291 certain background assumptions and in certain contexts). Philosophers, he  
 292 says, have 'hypostatized' parsimony. In other words, they have assumed that  
 293 appeals to parsimony mean the same thing in every context. Instead, he sug-  
 294 gests that when a scientist appeals to parsimony, his or her appeal has a distinct  
 295 meaning relative to a specific context and specific background assumptions.  
 296 Appeals to parsimony may thus be appeals to a theory's greater plausibility,  
 297 given either its greater likelihood, and prior probability, or a combination of  
 298 both. In his words, 'parsimony is a virtue that does not speak its name.' So,  
 299 appeals to parsimony in the assessment of competing hypotheses, may in fact  
 300 be appeals to something 'more fundamental' – namely, the greater likelihood,  
 301 or prior probability of some theory, relative to the evidence and our back-  
 302 ground beliefs. There is no reason to adopt Occam's razor, in the sense of  
 303 'fewest entities and/or processes' as a general principle for all of science all of  
 304 the time. Rather, the nature and relevance of considerations of parsimony are  
 305 context dependent. In Sober's terms, parsimony is not a 'global,' but a 'local,'  
 306 virtue of theories.

307 Sober's argument for this claim comes in two stages. First, he notes that if we  
 308 understand hypothesis evaluation using Bayesian framework, appeals to par-  
 309 simony make better sense of these arguments if we understand them in terms of  
 310 likelihoods and priors. Second, he argues that several cases of scientist's appeal  
 311 to parsimony are best interpreted in this way. Let's consider an example Sober  
 312 uses to illustrate his point. Williams argued that Wright's three phase shifting  
 313 balance process is not very likely; he then extends this argument to argue that  
 314 group-level selection in general is unlikely (1966, pp. 111–117).

315 Sober rationally reconstructs Williams' argument along Bayesian lines.<sup>6</sup> Let  
 316 us designate group level selection hypotheses as HG, and individual level  
 317 selection hypotheses as HI. First, Williams concedes that the phenomena (say,  
 318 a herd of fleet deer) could be equally well be a product of group-level selection  
 319 as individual-level selection. I.e. they have equal likelihoods.

<sup>6</sup> As my purposes here are expository, it is not relevant to this discussion whether Sober's reconstruction is faithful to the text. Indeed, Williams offers at least three (Sober and Wilson 1998) different rationales for the greater parsimony of lower-level selection hypotheses. The argument that moves from a critique of Wright's model to a critique of group selection is only one of several.



$$\Pr(e/HG) = \Pr(e/HI)$$

321 What he must mean, according to Sober, when he claims that lower level  
322 selection hypotheses are more parsimonious is that they have different priors,  
323 or  $\Pr(HG/e) < \Pr(HI/e)$ .

324 Williams argues that group selection requires a number of restrictive  
325 assumptions about population structure: there must be sufficient variation  
326 among groups, and rates of colonization and extinction must be sufficiently  
327 high. These facts, (which have since been contested) are the reasons he gives for  
328 the claim that lower level selection hypotheses are more parsimonious. I.e. they  
329 have higher priors, in light of the evidence concerning variation among groups  
330 and rates of colonization. According to Sober, Williams is not arguing that  
331 group selection never happens, he's simply suggesting that it's highly implau-  
332 sible that it does. According to Williams, the conditions for the possibility of  
333 group selection rarely hold, so:

$$\Pr(HI/e) > \Pr(HG/e) \text{ because } \Pr(HI) > \Pr(HG)$$

335 So, using the Bayesian framework, we can understand Williams's claims for  
336 the greater parsimoniousness of individual selection hypotheses as claims  
337 about the prior probability of individual selection being higher; where, prior  
338 probability here is simply understood as the probability, given our background  
339 beliefs about what conditions obtain in nature and which conditions are  
340 required for group vs. individual selection.

341 My (and Sober's) point here is not that Williams was correct.<sup>7</sup> Rather, the  
342 point is that this is one way of reconstructing scientists' appeals to parsimony  
343 that does not make them appear to be simple-minded (some humor). Some-  
344 times, of course, scientists make flawed arguments, and appeals to parsimony  
345 can be ad hoc justifications. Several of Williams arguments do fall under this  
346 category. However, his argument against Wright follows the same structure as  
347 that above, and the very same rationale is offered by Coyne et al. except with  
348 the advantage of 50 years of theoretical and empirical work on shifting bal-  
349 ance. Ultimately, whether or not a theory is understood to be more or less  
350 parsimonious may be understood as an argument over whether it has a higher  
351 likelihood, or a higher prior probability, or a combination of the two. In this  
352 case, the claim is simply that 'prior probability' is, in the Williams' context, a  
353 matter of plausibility of the conditions required for some process to occur. One  
354 might argue that it is dangerous to introduce formal tools when they simply  
355 cannot be defined suitably in these informal contexts, and that Sober does this  
356 to his detriment. It would be better, the objection goes, if Sober admitted that  
357 the Williams is not talking about probabilities *per se* but rather *judgments* of  
358 plausibility or truth. However, this is to miss the point of the exercise. Sober's  
359 claim is not that in fact scientists assign quantitative values to prior proba-

<sup>7</sup> Indeed, Sober has offered considerable argument to the contrary, see Sober and Wilson 1999.

bilities to this or that hypothesis, or that they in fact conditionalize using a Bayesian model. Rather, the point is that we may use a Bayesian framework to unpack the reasoning at work in judgments regarding competing hypotheses. And, if we do so, then appeals to parsimony do not seem to be merely matters of subjective opinion. Instead, they are judgments of plausibility based upon empirical data and background theory.

So, now let's turn to an analysis of the neo-Fisherians' argument and see whether this approach to parsimony can be helpful. When Coyne et al. say that Fisher's theory is more parsimonious than Wright's, what could they mean? First, they say that the presuppositions that Wright makes about the problem of adaptive evolution, those very ones which make shifting balance seem to provide a solution, are false. Second, they argue that the conditions required for all three phases in Wright's three phase process to operate in conjunction are not very often met in nature. Third, they claim that there is ample evidence that selection has generated adaptation, and, the conditions required for it to operate are not at all restrictive (All one requires is additive genetic variance, which they claim is amply available, and where  $4N_s > 1$ , selection must be the main factor in adaptive evolution).

Coyne et al.'s parsimony argument is not that Fisher's theory is preferable insofar as it invokes fewer mechanisms or processes. Rather, it is that the conditions required for the specific combination of mechanisms that makes possible Wright's model are unlikely to obtain. They are not suggesting that populations are not often subdivided, that drift plays no role in evolution, or that epistatic interactions between genes never occur. Rather, they are suggesting that the specific 'concatenation' of isolation, drift, intra- and inter-demic selection required by shifting balance does not often obtain, and so it is unlikely that Wright's model explains many (if any) adaptations in nature.

Coyne et al.'s argument is not simply an argument against the plausibility of shifting balance. It is additionally a prudential argument. In other words, given the implausibility of shifting balance, they claim that it is prudential to adopt selection as our working hypothesis, should we come across some apparent adaptation in our investigations. Ample empirical and theoretical evidence exists in favor of selection, so there is no question that it occurs. In contrast, while there is some evidence for at least two phases of Wright's three-phase process, the evidence in support of all three phases occurring in sequence is rather slim, or so they claim.

In my view, this is an instance where Sober's defense of a 'local' notion of parsimony coincides with the actual practice of appeals to parsimony in the sciences.<sup>8</sup>

---

<sup>8</sup> Note Coyne's comments (personal communication). He writes: I think you are correct in your interpretation... We say several times, I believe, that the SBT requires the concatenation of improbable circumstances AND that it is also untestable in many cases.

### 399 **Why Wright might not have been right**

400 Now, I'll discuss briefly some of the key points of Coyne et al.'s argument, and  
 401 then turn to issues of prudence in choice of research program in conclusion.  
 402 Their argument comes in three stages: first, they question the presuppositions  
 403 that led Wright to formulate his model. According to Coyne et al. the problem  
 404 as posed by Wright, of finding a 'trial and error' mechanism 'by which the  
 405 species may find its way from lower to higher peaks' was confused. Second,  
 406 they argue that each stage of the Shifting Balance individually is implausible,  
 407 for both empirical and theoretical reasons. Third, they claim that the condi-  
 408 tions required for all three stages to follow one upon another are restrictive and  
 409 unlikely to hold in nature.

410 What was the purported problem that Wright set out to solve? Wright was  
 411 motivated to develop the shifting balance theory in reply to what he saw as a  
 412 serious problem for adaptive evolution: the problem of escaping highly  
 413 adaptive 'peaks' in the field of gene combinations. Wright had been a student  
 414 of Castle at the Bussey institute; where the research program was focused on  
 415 the inheritance of traits due to multiple genes in interaction. Most of his  
 416 graduate work and his first job for the USDA were concerned with the  
 417 inheritance of complex traits, such as coat color, or milk yield. So, Wright was  
 418 impressed with the fact that many traits were influenced by multiple genetic  
 419 factors. Moreover, he was impressed with biogeographical work that seemed to  
 420 show that the differences between species did not appear to be adaptive  
 421 (Provine 1985 ■Au: AQ: Citation Provine (1985) is not in list please add to  
 422 list.■). He concluded from these observations that major transitions between  
 423 species could not possibly be the result of selection alone. The view that  
 424 selection alone was not sufficient to generate novel species is an old argument in  
 425 biology and one which found its way into textbooks popular at the time Wright  
 426 began his studies in biology (see, for instance, Kellogg 1903, one of the first  
 427 texts Wright read in his first courses in biology).

428 According to Wright, complex traits, in particular, adaptive traits, are most  
 429 likely the result of genes that are more or less fit in combination – i.e. that  
 430 there is pervasive epistasis for fitness. If there is pervasive epistasis for fitness,  
 431 then it is not possible for highly adaptive gene combinations to be broken up  
 432 without a population losing fitness. So, on this view, major adaptive changes  
 433 in a population require that a population traverse an 'adaptive valley' via  
 434 drift. Dobzhansky gives a vivid description of the adaptive landscape as  
 435 follows:

436  
 437 The field of gene combinations may, then, be visualized most simply in a  
 438 form of a topographical map, in which the contours symbolize the  
 439 adaptive values of various combinations. Groups of related combinations  
 440 of genes, which make the organisms that possess them able to occupy  
 441 certain ecological niches, are then, represented by the adaptive peaks  
 442 situated in different parts of the field. The unfavorable combinations of

443 genes which make their carriers unfit to live in any existing environment  
 444 are represented by the 'adaptive valleys' which lie between the peaks  
 445 (Dobzhansky 1951, pp. 8–9 ■Au: AQ: Citation Dobzhansky (1951) is not  
 446 in list please add to list.■).

447 According to Wright, it is necessary for a population's complex of genes to  
 448 be altered by sampling, or drift, in order for it to move to a more highly  
 449 adapted state. Since adaptation is a product of genes in combination, 'novel  
 450 gene combinations' are necessary for novel adaptations. If simply under the  
 451 control of selection, a species will ultimately come to rest on a suboptimal  
 452 peak. Notice that the above argument is an inference from an number of  
 453 observations to a rather wide-ranging conclusion about adaptation and evo-  
 454 lution as a whole.

455 Here's Wright's statement of the problem that he sought to solve with the  
 456 shifting balance theory:

457 458 The problem of evolution as I see it is that of a mechanism by which the  
 459 species may continually find its way from lower to higher peaks. ... in  
 460 order that this may occur, there must be some trial and error mechanism  
 461 on a grand scale by which the species may explore the regions sur-  
 462 rounding the small portion of the field which it occupies. To evolve, the  
 463 species must not be under the strict control of selection (Wright 1932, pp.  
 464 163–164 ■Au: AQ: Citation Wright (1932) is not in list please add to  
 465 list.■).

466 Wright's trial and error mechanism was drawn directly from the breeding  
 467 program of the Duchess Shorthorn cattle, which he spent five years studying at  
 468 the USDA. In order to improve the cattle stock, the following procedure was  
 469 employed:

470 471 The first step in any case should be selection of a vigorous foundation,  
 472 approaching as closely as possible to the desired type. With such a  
 473 foundation stock, one might practice the most intensive inbreeding in a  
 474 large number of distinct lines, knowing that most lines would inevitably  
 475 deteriorate greatly, but trusting that a few would be found in which  
 476 desirable qualities would become fixed, and in which the deterioration in  
 477 any vital respect would be so slight that they could be maintained suc-  
 478 cessfully. By crossing such lines which have withstood this acid test of  
 479 inbreeding, one might reasonably hope to recover more than the original  
 480 vigor and retain those characters which had been fixed... This method, an  
 481 alternation of intensive inbreeding with selection and crossbreeding of the  
 482 few successful lines must naturally be done on a large scale... It is an  
 483 important method and has some parallel in the general history of the  
 484 breeds. Many of the early breeders practiced [it]. (Wright (1923b) ■Au:  
 485 AQ: Cited reference Wright (1923b) is not in list please add to list.■, in  
 486 Provine (1986), p. 46.)

487 Here, in a 1923 discussion of cattle breeding, Wright gave a preliminary  
 488 statement of what would become his three phase 'shifting balance' model of  
 489 evolution:

490 Phase I: Genetic drift causes local populations to temporarily lose fitness,  
 492 shifting across adaptive valleys toward new, higher adaptive peaks.

493 Phase II: Selection within demes places them atop new peaks.

495 Phase III: Different adaptive peaks compete with one another, causing  
 497 fitter peaks to spread through the entire population. Or, migration out  
 498 from the most adaptive deme leads to the spread of the most adaptive.

499 In sum, given the pervasiveness of epistasis for fitness, which Wright wit-  
 500 nessed in his experimental and USDA work, if strictly under the control of  
 501 selection, a population could not make significant adaptive changes. (Note here  
 502 that he's moving from the case of multi-genic traits in mammals to all traits in  
 503 all species.) So, in order to escape suboptimal gene combinations, a population  
 504 needs to be broken up into small subpopulations, which, after a period of  
 505 isolation (during which drift and intrademic selection enable a population to  
 506 'escape' undesirable gene combinations), can then come into contact and  
 507 compete.

508 Coyne et al. question the problem of evolution as set by Wright. Or, they  
 509 question Wright's rationale as to why one must invoke explanations other than  
 510 selection for adaptive evolution. Wright supposed that mass selection is too  
 511 slow to explain diversity, that mutation is insufficient as source of variation,  
 512 that cost of substitutions constrains the rate of adaptation, and that phenotypic  
 513 change involves the appearance of maladaptive intermediates. In other words,  
 514 he thought that the Darwinian paradigm, according to which selection acts in  
 515 relatively large, panmictic populations, with mutations as the 'raw material',  
 516 was not adequate to account for complex adaptations and the diversity of life.  
 517 Some new story needed to be told that will explain how it is possible that new,  
 518 more adaptive, gene combinations can come about.

519 Coyne et al. deny that all of these are legitimate worries. I'll focus on the last  
 520 and one of the most longstanding objections: that phenotypic change involves  
 521 the appearance of maladaptive intermediates, since this is what fueled Wright's  
 522 idea of the adaptive landscape, and what ultimately lead to his shifting balance  
 523 idea.

524 Wright's claim is that given the extent of epistatic interaction for fitness, we  
 525 require some mechanism to explain how it is that a population can move from  
 526 one highly adapted gene combination to another more adaptive combination.  
 527 Crow (1990) ■Au: AQ: Citation Crow (1990) is not in list please add to list.■  
 528 describes the situation using a simple haploid model as follows:

529 Suppose alleles a and b go well together, as do A and B, but A and b  
 531 and a and B do not. Suppose further that the AB combination is better

than the ab. If a population has a high frequency of a and b alleles it will not move to a state in which A and B are common, because to do so will produce a large number of inferior Ab and aB recombinants. We can think of this as a three-dimensional graph in which the two abscissas are the frequencies of the Ab and Ba alleles and the ordinate is the mean fitness of a population with this frequency combination. The surface will be saddle shaped, with a low peak where ab is common and a high one where AB is common. A population near the lower peak cannot get to the higher one without crossing a valley of lower fitness. (Crow 1990, p. 75)

Coyne et al. reply to this argument as follows. Wright imagines that the adaptive landscape is static, or, that the mean fitness of a population will be constant. However, there is good reason to think that this is false. First, a population may shift into an adaptive valley for any number of reasons other than drift – change in environment, for instance. A change in environment or a mutational change could change the mean fitness of a population, such that a population originally on a peak may come to rest on a valley, and simple selection could pull it up to a new peak. Moreover, ‘ridges’ can arise between adaptive peaks for one of several reasons. The fitness of a particular gene combination changes due to its relative frequency in a population, or because of the relative numbers of other individuals in the same environment. This phenomenon, that the fitness of a particular trait can change because of the relative frequency of individuals possessing this trait, is known as frequency dependence. Thus, particular genotypic combinations are not necessarily ‘stuck’ atop adaptive peaks. Indeed, phenomena like frequency dependence challenge the whole idea of a three dimensional landscape. If different individuals are more or less fit relative to the number other kinds of individuals in their cohort, then it does not make sense to speak of specific genotypes having specific fixed fitnesses. Not only will the adaptive landscape will be constantly changing because of the selection coefficient of some trait will change with frequency dependence, but if we consider the many dimensions in which we can measure an individual’s (or a population’s!) fitness, there are multiple ridges that an individual (or population) may traverse via selection. So, Coyne et al.’s first argument against Wright is that the problem of evolution as he describes it is not the problem he imagined.

Their second line of attack is to suggest that individually and in combination, each phase of the shifting balance process is unlikely to occur in nature, on both theoretical and empirical grounds. It is true that the chance of peak shifting by drift increases with decrease in population size. However, the chances of staying atop an adaptive peak for very small populations is very small. In other words, the smaller one’s reserve of variation, or, what is the same, the smaller the population size, the more likely that a population will drift into a valley or simply die out than that it will drift toward the vicinity of a new, more adaptive peak. This renders phase I implausible.

576 Coyne et al. further point out that many processes besides drift that can  
 577 move populations to different peaks (phase II). As mentioned above, local  
 578 peaks could be converted to ridgesa, allowing adaptive advance by selection.  
 579 And, of course, every case of natural selection is just a case of phase II.

580 Second, with respect to Phase III: peak shifts may occur only in sparsely  
 581 populated parts of a species' range (i.e. the subpopulations would have to be  
 582 so isolated as to have none or very little incoming variation); so it's difficult  
 583 to see how novel peaks could spread, or, how more highly adapted groups  
 584 could come into competition with other groups. Third, with respect to  
 585 empirical evidence: they claim that for each example canvassed, phase I is  
 586 infrequent, and only one known case shows convincing evidence for phase  
 587 III. In sum, they suggest that theory shows that shifting balance can some-  
 588 times be an efficient mechanism in adaptive evolution, but only under  
 589 restrictive conditions. And, empirical evidence suggests that there are very  
 590 few cases, if any, where all three phases in sequence have actually occurred.  
 591 Fisherian mass selection process is thus, they claim, "more parsimonious"  
 592 than the shifting balance process.

### 593 Conclusions

594 What scientists mean by parsimony will vary relative to context, and, these  
 595 different senses of parsimony can yield more or less epistemically sound  
 596 grounds for adopting one or another theory or research program. Further,  
 597 parsimony in the context of the Wright–Fisher debate is *not* greater simplicity  
 598 in the sense of economy of process.<sup>9</sup> Rather, parsimony in this case amounts to  
 599 plausibility, in this case, there is substantial evidence that Wright's presuppo-  
 600 sitions are false, and that his mechanism operates rarely in nature. In light of  
 601 this interpretation of parsimony, I think that we can understand why Coyne et  
 602 al. think we ought to be skeptical of the significance of shifting balance. They  
 603 are offering an argument from what one might call prudence. Or, perhaps  
 604 better, given the empirical evidence, we ought to be cautious as to whether  
 605 shifting balance has played an important role in the evolutionary process.  
 606 Caution or prudence are not words commonly used in scientific contexts; more  
 607 often, scientist will appeal to a vague notion such as 'parsimony', which  
 608 inevitably leads to confusion and conflation.

609 This is not to deny that there are many mechanisms at work in nature. The  
 610 claim is not that there is no population structure, drift, or epistatic interactions  
 611 between genes. All they are suggesting is that unless and until we are called  
 612 upon to do otherwise, it's prudent to adopt the Fisherian model in attempting

---

<sup>9</sup> Though, of course, it is an open question what exactly we mean by 'economy of process'. Does the more economical process contain fewest kinds, fewest new kinds, or simply the fewest number of entities or processes? (Thanks to Marc Lange for this comment).

to explain some adaptation. The authors of these papers agree that there are a multiplicity of mechanisms at work in evolution. They agree that selection, drift, epistasis, etc. are all important factors in the process of evolution. What they disagree about is how often it is the case that one particular combination of these mechanisms obtains: namely, Wright's three phase shifting balance process.

Were Coyne et al. conflating explanatory power, generality and parsimony? No. Adopting Coyne et al.'s view does not require of us that we rule out drift, epistasis, or population structure as important factors in explaining evolutionary pattern or processes. We may still be pluralists about the many mechanisms and processes at work in generating evolutionary change. All it requires is being suspicious of one particular combination of these factors. Coyne et al. are not suggesting that mass selection is a *sufficient* explanation for adaptation, only that it is the a very likely candidate, barring evidence for isolation, drift, etc.

A nice analogy to this suggestion comes from Mayr's (and later, Sober's) discussion of the adaptationist program. We can separate the question of whether we ought to adopt adaptationism as a research program from the question of whether natural selection is sufficient to explain any or all particular adaptations. Likewise, we may separate the question of whether we ought to adopt a Wrightian research program from whether in fact Wright's model explains any particular adaptation. There may be good reasons to reject the research program, but this is not to say that Wright's three phase process never occurs.

The kind of arguments I think that one might sensibly offer in support of adopting a research program will be very different from those offered in support of invoking a particular mechanism to explain some particular phenomena. Plausibility arguments of the sort I've just discussed will figure in the former discussions, but not in the latter. For any particular case, however, if the evidence isn't decisive, it's not clear that we much adopt one or the other. People can maintain two or three alternatives. Feynman has written that all the good physicists he knew kept about 3 models in their head all the time, and interpreted new evidence on each of them. Analogously, keeping in mind the roles of drift, mutation, migration, and various forms of selection (frequency dependent, intra- and inter-demic) is a good strategy. I think that we may consistently say that Fisherian mass selection is an important explanation for adaptation, without also committing ourselves to the view that mass selection is a sufficient explanation – or, that it suffices to explain every particular adaptation.

## Acknowledgements

Thanks to Steve Downes, Marc Lange, Jay Odenbaugh, Paul Sniegowski, Charles Twardy and an anonymous reviewer for their thoughtful comments and suggestions.



## 655 References

- 656 Carlson 1966. *The Gene: A Critical History*.  
 657 Coyne, Barton and Turelli 1997. Perspective: a critique of Wright's shifting balance theory of  
 658 evolution. *Evolution* 51(3): 643–671.  
 659 Dobzhansky 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.  
 660 \*Later editions: 1951.  
 661 Gavrillets 1996. On phase three of the shifting balance theory. *Evolution* 50(3): 1034–1041.  
 662 Gavrillets 1997. Evolution and speciation on holey adaptive landscapes. *Trend. Ecol. Evol.* 12(8):  
 663 307–312.  
 664 Gould and Lewontin 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique  
 665 of the Adaptationist Programme. *Proc. Roy. Soc. Lond. B Biol. Sci. Evol. Adapt. Nat. Sel.* 205,  
 666 (1161): 581–598.  
 667 Jacob F. 1977. Evolution and tinkering. *Science* 196: 1161–1166.  
 668 Kellogg 1903. *Darwinism To-day: a Discussion of Present-Day Scientific selection Theories*. Bell,  
 669 London.  
 670 Lewontin 1999. What do population geneticists know and how do they know it?. In: Creath and  
 671 Maienschein (eds), *Biology and Epistemology*. Cambridge University Press, Cambridge.  
 672 Ockham 1957. *Philosophical Writings; A Selection*. edited and translated by Philotheus Boehner.  
 673 Nelson, Edinburgh, New York.  
 674 Newton 1729. *The Mathematical Principles of Natural Philosophy*, trans. Andrew Motte, vol. 2  
 675 (London: Printed for B. Motte), pp.202–205.  
 676 Ruse 1993. Are Pictures Really Necessary? The Case of Sewall Wright's 'Adaptive Landscapes' (in  
 677 *Biology: The Non-Propositional Side*). *PSA: Proceedings of the Biennial Meeting of the Phi-*  
 678 *losophy of Science Association*. 1990, Volume Two: Symposia and Invited Papers. 1990: 63–77.  
 679 Sober and Elliott 1990. Let's Razor Ockham's Razor. *Philosophy: J. Roy. Inst. Phil.* 1990(Supp):  
 680 73–93.  
 681 Sober and Elliott 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. MIT Press,  
 682 Cambridge.  
 683 Sober and Wilson 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*.  
 684 Harvard University Press, Cambridge MA.  
 685 Skipper R. 2002. The Persistence of the RA. *Fisher-Sewall Wright Controversy. Biol. Phil. Je* 02;  
 686 17(3): 341–367.  
 687 Wade, Michael, Goodnight and Charles J. . Perspective: theories of Fisher and Wright in the  
 688 context of metapopulations: when nature does many small experiments. *Evolution* 52(6): 1537–  
 689 1553 (Dec. 1998). ■Au: AQ: Except reference no. 7 & 16 please provide the Initials for the  
 690 authors in all the references.■